H^2GT: Hub-enhanced Hierarchical Graph Transformer for Explainable Diagnosis of ASD with Functional Brain Networks

Yiqi Xiu[1], Geng Chen[1(✉)], Yin Huang[1], Xinhong Li[1], Xuyun Wen[2], and Shijie Zhao[3]

[1] National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, China
[2] College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[3] School of Automation, Northwestern Polytechnical University, Xi'an, China
✉ geng.chen@ieee.org

Abstract. Autism Spectrum Disorder (ASD) is a neurodevelopmental condition marked by social and behavioral abnormalities. Functional brain networks (FBNs) are widely used in ASD diagnosis. While modular organization and hub nodes are key topological features of FBNs, existing methods fail to leverage them effectively for explainable diagnosis or to jointly model intra-module and inter-module interactions. We propose a Hub-enhanced Hierarchical Graph Transformer (H^2GT) to address these issues. It integrates hub-guided topological priors into a multi-scale graph Transformer, enabling joint learning of local node interactions and global module coordination. Specifically, a Hub-Prioritized Attention Mechanism highlights hub nodes, while dedicated components model fine-grained intra-module connectivity and cross-module coordination. Experiments on ABIDE I and II show that H^2GT achieves state-of-the-art diagnostic performance and identifies multi-scale biomarkers, offering deeper insights into ASD pathology.

Keywords: Autism Spectrum Disorder · Functional brain networks · Graph Transformer · Hub nodes · Interpretable diagnosis.

1 Introduction

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by persistent deficits in social communication and interaction, alongside restricted and repetitive behavioral patterns. Clinical diagnosis remains challenging due to symptomatic heterogeneity and the absence of definitive biological markers, relying primarily on behavioral assessments [12]. Neuroimaging technologies, particularly functional Magnetic Resonance Imaging (fMRI), offer a promising avenue by enabling the construction of Functional Brain Networks (FBNs). These networks reveal significant alterations in functional connectivity (FC) in key regions such as the prefrontal cortex and temporal lobes

in individuals with ASD, providing critical insights into its pathological mechanisms [27,24].

Deep learning methods have demonstrated significant promise in analyzing complex brain networks [18,33,6,14]. BrainNetCNN [18] employs specialized convolutional filters tailored for brain network analysis. Graph Neural Networks (GNNs) [33], including BrainGNN [19], effectively model local functional relationships for biomarker identification, while FBNETGEN [15] provides a task-aware interpretable framework. Graph Transformer [10] excels at capturing global, long-range dependencies and integrating graph-specific processing, yet this architecture is not specifically designed for brain networks. The Brain Network Transformer (BNT) [16] adapts the Transformer architecture to brain graphs, enabling global dependency modeling without relying on predefined graph structures.

However, these prevailing architectures face inherent limitations. GNNs struggle to model system-level coordination across functional modules due to their localized message-passing schemes [19], while Transformers often overlook fine-grained, within-module topological interactions [16]. Simultaneously capturing both local interactions and global coordination, fundamental operating principles of the brain, remains a significant challenge. Furthermore, most methods prioritize diagnostic performance over clinical interpretability, failing to integrate well-established topological priors of FBNs, such as modular organization and hub nodes [30], which are crucial for functional integration and differentiation [20]. The brain exhibits distinct modular organization [32,28,29] with hub nodes playing critical roles [31,26], classified as provincial and connector hubs [21,2]. ASD research has identified atypical patterns in key networks such as the Default Mode Network (DMN) [22] and Dorsal Attention Network (DAN) [11], as well as hubs [27,3]. However, existing methods rarely identify these critical biological entities, limiting traceable biological evidence and clinical utility [25,5].

To address these challenges, we propose Hub-enhanced Hierarchical Graph Transformer (H^2GT), an explainable model for ASD diagnosis with FBNs. This framework integrates topological priors into a multi-scale graph Transformer, jointly modeling fine-grained intra-module interactions and global inter-module coordination. It incorporates a biologically grounded hierarchical design that captures both local and global dependencies. Identified hub nodes are encoded as spatial attention biases to preserve critical brain connectivity patterns, and a topology-guided mechanism simulates brain-like information flow to enhance interpretability. Experiments on the ABIDE I and II dataset [4,9] demonstrate state-of-the-art diagnostic performance and reveal multi-scale biomarkers at the node, hub, and module levels, offering deeper insights into ASD pathology.

2 Method

2.1 Overview

As shown in Fig. 1, the proposed H^2GT operates on FBNs derived from fMRI data. First, the brain undergoes modular partitioning and topological rearrange-

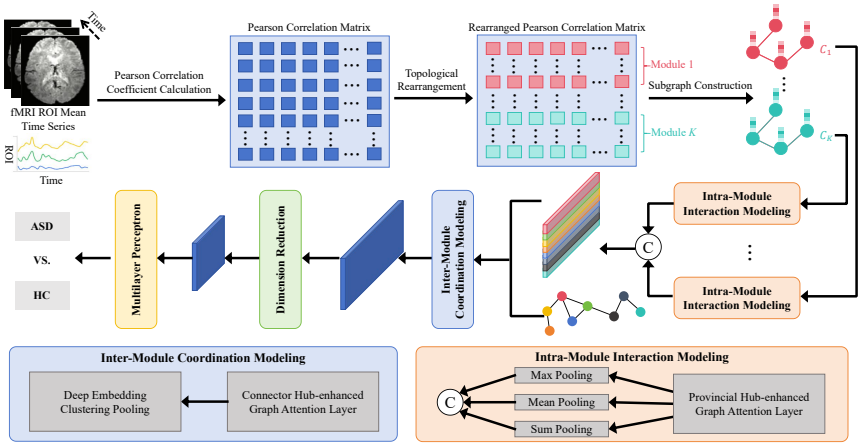


Fig. 1: Schematic overview of the proposed H^{2}GT framework.

ment of regions of interest (ROIs) to create hierarchical functional subnetworks. Subsequently, Intra-module Interaction Modeling extracts features within modules, while Inter-module Coordination Modeling captures interaction patterns across modules. Finally, classification is performed based on whole-brain representations.

2.2 Hierarchical Graph Construction

For a single subject, we construct a brain graph $G = (\mathbf{A}, \mathbf{X})$, where $\mathbf{A} \in \{0,1\}^{M\times M}$ is the adjacency matrix and $\mathbf{X} \in \mathbb{R}^{M\times M}$ is the node feature matrix. M denotes the number of ROIs. The adjacency matrix is derived by thresholding the absolute Pearson correlation matrix $\boldsymbol{\rho} \in [-1,1]^{M\times M}$ computed from fMRI time series, retaining only the strongest connections. The feature vector of node v is its full correlation profile: $\mathbf{X}(v) = [\rho_{v1}, \rho_{v2}, \ldots, \rho_{vM}]^\top$.

To capture the hierarchical organization of the brain, we partition the ROIs into K functional modules according to a predefined atlas. This yields a set of subgraphs $\{G^1, G^2, \ldots, G^K\}$, where each subgraph $G^k = (\mathbf{A}^k, \mathbf{X}^k)$ contains the nodes belonging to module k and the corresponding induced edges and features.

We further construct a module-level graph $G^{\text{mod}} = (\mathbf{A}^{\text{mod}}, \mathbf{H})$. Each node in G^{mod} corresponds to one functional module; its node feature $\mathbf{h}_k \in \mathbf{H}$ is obtained by processing the subgraph G^k with a graph Transformer layer (Section 2.4). The adjacency matrix $\mathbf{A}^{\text{mod}} \in \mathbb{R}^{K\times K}$ represents cross-module connectivity: $A^{\text{mod}}_{jk} = \frac{1}{|j||k|}\sum_{u\in j, v\in k} |\rho_{uv}|$, where $|j|$ and $|k|$ are the sizes of modules j and k, respectively.

2.3 Hub Node Detection and Prioritization

Hub Detection. To guide the model's attention toward neurobiologically critical nodes, we identify hub nodes based on a group-level FC matrix estimated via DGN [13]. For a given node, its within-module degree z-score (WD) and participation coefficient (PC) are defined as:

$$\mathrm{WD} = \frac{e^m - \bar{e}^m}{\sigma_e^m}, \qquad \mathrm{PC} = 1 - \sum_{c=1}^{K} \left(\frac{e^c}{d}\right)^2, \tag{1}$$

where m is the module containing the node, e^c denotes the number of connections from the node to nodes in module c, $d = \sum_{c=1}^{K} e^c$ is the total degree of the node, and $\bar{e}^m$, σ_e^m represent the mean and standard deviation of intra-module degrees across nodes within module m. Nodes are then classified as *provincial hubs* (high WD, low PC) or *connector hubs* (high WD, high PC). Let $\mathcal{H}_{\mathrm{prov}}$ and $\mathcal{H}_{\mathrm{conn}}$ denote the sets of provincial and connector hubs, respectively.

Hub-Prioritized Attention Mechanism. We incorporate this topological prior by augmenting the standard self-attention with a learnable hub bias matrix $\mathbf{B}_{\mathrm{hub}} \in \mathbb{R}^{M \times M}$:

$$\mathrm{Attention}(Q, K, V) = \mathrm{Softmax}\left(\frac{QK^\top}{\sqrt{d_k}} + \lambda \cdot \mathbf{B}_{\mathrm{hub}}\right) V, \tag{2}$$

where Q, K, V are the query, key, and value matrices, d_k is the dimension of the key vectors, and λ is a coefficient controlling the strength of the prior. The bias term is defined as:

$$\mathbf{B}_{\mathrm{hub}}[i, j] = \begin{cases} \beta_{\mathrm{prov}}, & \text{if } j \in \mathcal{H}_{\mathrm{prov}}, \\ \beta_{\mathrm{conn}}, & \text{if } j \in \mathcal{H}_{\mathrm{conn}}, \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

where β_{prov} and β_{conn} are trainable parameters. This mechanism amplifies attention weights toward hub nodes, mimicking their central role in biological information integration.

2.4 Multi-scale Feature Learning with Graph Transformers

Intra-Module Modeling with Provincial Hub Enhancement. Each subgraph G^k is processed by a Provincial Hub-enhanced Graph Attention (PHGA) layer, which applies the hub-prioritized attention mechanism to model fine-grained interactions within the module:

$$\mathbf{H}^k = \mathrm{PHGA}\big(G^k; \theta^k\big) \in \mathbb{R}^{n_k \times d}, \tag{4}$$

where n_k is the number of nodes in module k, d is the hidden dimension, and θ^k denotes the layer parameters. A combination of max, mean, and sum pooling operation aggregates node features into a compact module embedding $\tilde{\mathbf{h}}_k \in \mathbb{R}^{3d}$:

$$\tilde{\mathbf{h}}_k = \big[\max(\mathbf{H}^k) \parallel \text{mean}(\mathbf{H}^k) \parallel \text{sum}(\mathbf{H}^k)\big], \tag{5}$$

where $\parallel$ denotes concatenation, and max, mean, sum are applied row-wise.

Inter-Module Modeling with Connector Hub Enhancement. The module embeddings $\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_K$ form the node feature matrix $\mathbf{H} \in \mathbb{R}^{K \times 3d}$ of the module-level graph G^{mod}. A Connector Hub-enhanced Graph Attention (CHGA) layer, which also employs the hub bias matrix $\mathbf{B}_{\text{hub}}$ (now focusing on connector hubs), models global coordination patterns across modules:

$$\mathbf{Z} = \text{CHGA}\big(G^{\text{mod}}; \phi\big) \in \mathbb{R}^{K \times d'}. \tag{6}$$

The output $\mathbf{Z}$ is then pooled via a Deep Embedding Clustering (DEC) mechanism, which learns a cluster assignment matrix to aggregate module-level representations [17]. The resulting features are flattened and fed into a multi-layer perceptron (MLP) to produce the final diagnostic prediction $\hat{y}$.

2.5 Optimization Objective

The entire H^{2}GT is trained end-to-end using the binary cross-entropy loss:

$$\mathcal{L} = -\frac{1}{B}\sum_{i=1}^{B}\big[y_i \log \sigma(\hat{y}_i) + (1 - y_i)\log(1 - \sigma(\hat{y}_i))\big], \tag{7}$$

where B is the batch size, y_i is the ground-truth label for sample i, $\hat{y}_i$ is the corresponding model output, and $\sigma(\cdot)$ is the sigmoid function.

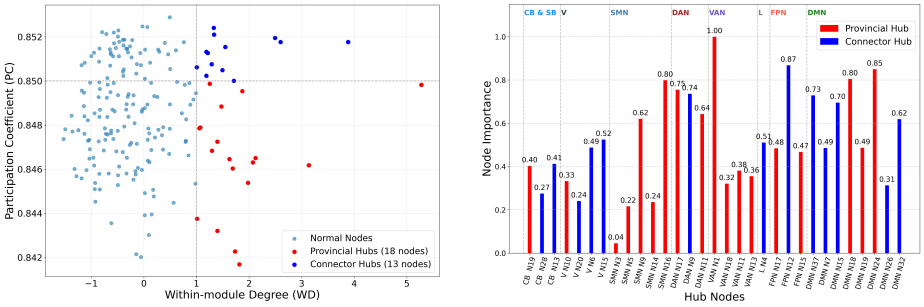
3 Experiments

3.1 Data Description and Preprocessing

The H^{2}GT framework was evaluated on ABIDE I [4] and an ABIDE II-derived dataset [9].ABIDE I includes resting-state fMRI of 493 ASD and 512 healthy controls (HC), preprocessed with CPAC [7]. Functional connectivity matrices were constructed by Pearson correlation of mean time series from 200 Craddock atlas regions [8], assigned to eight Yeo-7 functional modules [34]. The ABIDE II-derived dataset contains 344 ASD and 402 HC, identically preprocessed, with AAL atlas (116 ROIs) connectivity matrices assigned to six common functional modules.

Table 1: Performance comparison on ABIDE I and II datasets (%).

Model	ABIDE I				ABIDE II			
	AUC	Acc	Sen	Spe	AUC	Acc	Sen	Spe
MLP	50.5	51.8	59.4	44.7	58.0	61.0	52.8	68.3
GCN	69.8	61.6	74.2	49.9	59.8	62.3	58.3	65.9
GIN	62.4	59.1	40.6	76.3	63.9	61.0	66.7	56.1
GATE	59.7	59.8	64.5	54.9	62.0	62.3	65.8	58.2
BrainGNN	65.7	61.8	74.5	47.4	68.0	64.0	67.6	60.8
FBNETGNN	76.0	68.8	68.6	69.0	65.0	61.7	61.9	64.4
BrainNetCNN	79.7	71.9	74.3	69.0	64.2	60.0	57.1	64.3
BNT	77.8	71.9	71.9	**71.9**	68.7	64.3	56.3	**71.1**
Com-BrainTF	79.6	72.5	80.1	65.7	74.0	66.2	65.8	66.7
H^2GT	**82.2**	**75.0**	**81.3**	68.8	**76.5**	**72.4**	**75.8**	68.0



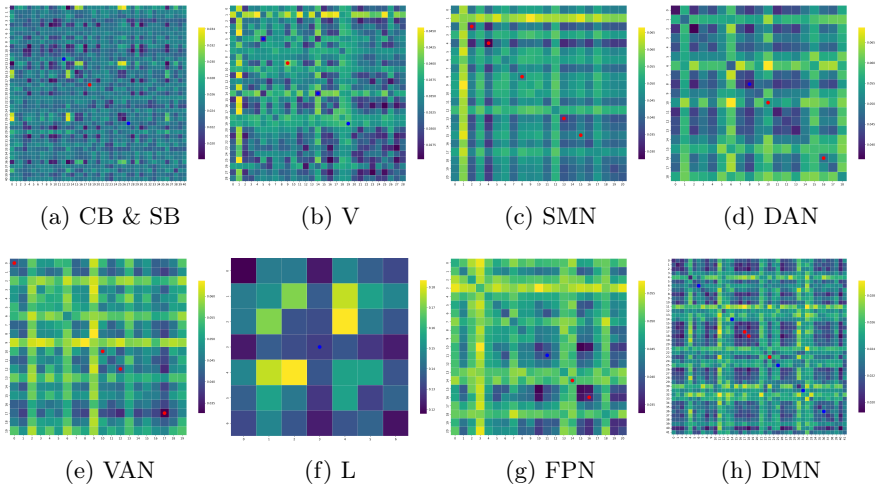
(a) Hub node detection results

(b) Hub node importance

Fig. 2: Visualized results of hub node analysis.

3.2 Experimental Setup

All models are implemented in PyTorch and trained on an NVIDIA 3060 GPU with 12GB memory. For both local and global graph Transformers, the number of attention heads is set equal to the number of modules. The Adam optimizer is used with an initial learning rate of 10^{-4} and a weight decay of 10^{-4}. The data is divided into train/validation/test sets with a ratio of 7:1:2, and a batch size of 64 is used. The ASD vs. HC prediction is a binary classification task. The performance on the test set is evaluated using AUROC, accuracy, sensitivity, and specificity. Models are trained for 50 epochs, and early stopping is employed. The epoch with the highest AUROC on the validation set is selected for comparison with the test set performance. Our code is publicly available at `https://github.com/xiuyiqi/H2GT`.



(a) CB & SB (b) V (c) SMN (d) DAN

(e) VAN (f) L (g) FPN (h) DMN

Fig. 3: Attention heatmaps of the eight functional modules. Red dots indicate provincial hubs and blue dots indicate connector hubs.

3.3 Comparison with Baseline Models

We evaluated H^{2}GT against nine baselines spanning general-purpose models (MLP, GCN, GIN), brain-network-specific models (GATE [23], BrainGNN [19], FBNETGNN [15], BrainNetCNN [18]), and Transformer-based architectures (BNT [16], Com-BrainTF [1]). As shown in Table 1, on ABIDE I and II, H^{2}GT achieves the highest AUC, accuracy, and sensitivity with competitive specificity. The consistent performance across both datasets demonstrates the robustness and effectiveness of our approach. This advantage stems from explicitly modeling hierarchical, hub-prioritized brain organization to capture both fine-grained intra-module interactions and global inter-module coordination.

3.4 Interpretability Analysis

A key strength of H^{2}GT is its inherent multi-scale interpretability, which yields biomarkers at the hub, ROI, and module levels, offering coherent biological insights.

Hub Node Analysis. As visualized in Fig. 2, the learned attention biases successfully identify neurobiologically critical provincial and connector hub nodes. Within each functional module, most of these hubs receive high attention weights (marked by red and blue dots in the Fig. 3), confirming their pivotal role in information integration for ASD classification. A small subset shows lower importance, likely reflecting that while structurally central, their functional relevance to ASD-specific activity patterns may be limited.

Critical ROI Analysis. Fig. 3 presents the attention scores for ROIs across the eight functional modules. The heatmaps reveal discriminative ROIs in modules such as the DAN and DMN, where many regions exhibit high attention

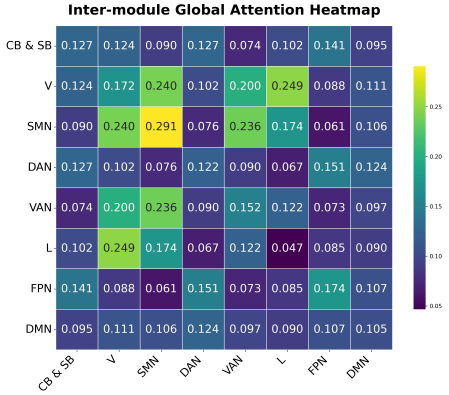


Fig. 4: Inter-module global attention heatmap.

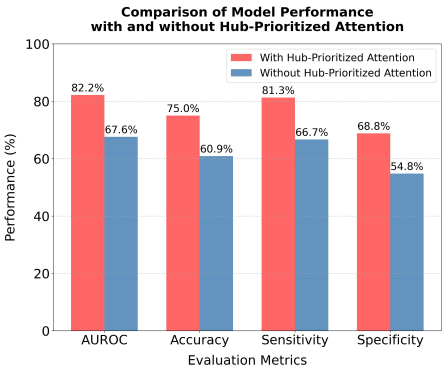


Fig. 5: Comparison of H^{2}GT with and without Hub-Prioritized Attention.

scores. For example, in the DAN module, several nodes show consistently elevated attentions, aligning with prior reports of atypical DAN connectivity in ASD [11]. Similarly, the DMN module contains clusters of high-attention ROIs, consistent with the well-documented DMN abnormalities in ASD [22]. The two categories of hub nodes are also marked in each internal heatmap, highlighting their spatial distribution.

Critical Functional Module Analysis. At the coarsest scale, Fig. 4 reveals the strongest interactions among several modules. Notably, the SMN and V modules exhibit high intra-module attention on their diagonals, indicating strong internal information processing. More importantly, L shows remarkably high cross-module attention with the V and SMN modules, suggesting that the L module integrates extensively with sensory and motor circuits. This finding is strongly supported by prior studies highlighting profound limbic system abnormalities in ASD [27]. The consistency between the model-identified biomarkers and established medical evidence underscores the potential clinical utility of H^{2}GT.

3.5 Ablation Experiments

This section rigorously validates the effectiveness of the Hub-Prioritized Attention Mechanism. Fig. 5 presents a detailed comparison of performance metrics for H^{2}GT both with and without the inclusion of Hub-Prioritized Attention. The results clearly demonstrate that models incorporating this attention mechanism consistently achieve superior performance across multiple evaluation criteria. This improvement emphasizes the importance of prioritizing hub regions within the brain network, as it enhances the model's ability to capture critical functional interactions, which are crucial for accurate prediction. Therefore, the Hub-Prioritized Attention module plays a pivotal role in enhancing the overall effectiveness of H^{2}GT.

4 Conclusions

In this work, we propose H^2GT, an FBN-based framework addressing incomplete multi-scale modeling and limited explainability in ASD diagnosis. Its innovations include a hub-prioritized attention mechanism, a hierarchical architecture capturing local-global brain interactions, and multi-scale interpretability, thereby enabling biologically grounded multi-scale modeling. Experiments show H^2GT outperforms existing methods. Ablation confirms the hub-prioritized attention's contribution, and explainability identifies ASD-relevant hubs, regions, and modules, offering biological insights. This advances feature extraction and interpretability, enhancing clinical utility. Future directions include dynamic hubs, multimodal data, extension to other disorders, and efficiency for clinical translation.

Acknowledgments. This study was funded by in part by the National Natural Science Foundation of China (62571451, U23A20335, 62522315, and 42271315), the Innovation Capability Support Program of Shaanxi (2025ZC-KJXX-47), and the Science and Technology Support Project of Guizhou Province ([2025]131).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bannadabhavi, A., Lee, S., Deng, W., Ying, R., Li, X.: Community-aware transformer for autism prediction in fmri connectome. In: International conference on medical image computing and computer-assisted intervention. pp. 287–297. Springer (2023)
2. Bertolero, M.A., Yeo, B.T., D'Esposito, M.: The modular and integrative functional architecture of the human brain. Proceedings of the National Academy of Sciences **112**(49), E6798–E6807 (2015)
3. Bullmore, E., Sporns, O.: The economy of brain network organization. Nature reviews neuroscience **13**(5), 336–349 (2012)
4. Cameron, C., Yassine, B., Carlton, C., Francois, C., Alan, E., András, J., Budhachandra, K., John, L., Qingyang, L., Michael, M., et al.: The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. Frontiers in Neuroinformatics **7** (2013)
5. Canario, E., Chen, D., Biswal, B.: A review of resting-state fmri and its use to examine psychiatric disorders. Psychoradiology **1**(1), 42–53 (2021)
6. Chen, G., Wang, Q., Feng, Y., Rekik, I.: Metadata-driven federated learning of connectional brain templates in non-iid multi-domain scenarios. IEEE Transactions on Medical Imaging **45**(4), 1513–1525 (2026)
7. Craddock, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B.S., Lewis, J.D., Li, Q., Milham, M., et al.: The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. Frontiers in Neuroinformatics **7**(27), 5 (2013)

8. Craddock, R.C., James, G.A., Holtzheimer III, P.E., Hu, X.P., Mayberg, H.S.: A whole brain fmri atlas generated via spatially constrained spectral clustering. Human brain mapping **33**(8), 1914–1928 (2012)
9. Di Martino, A., O'connor, D., Chen, B., Alaerts, K., Anderson, J.S., Assaf, M., Balsters, J.H., Baxter, L., Beggiato, A., Bernaerts, S., et al.: Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii. Scientific data **4**(1), 1–15 (2017)
10. Dwivedi, V.P., Bresson, X.: A generalization of transformer networks to graphs. In: AAAI Workshop on Deep Learning on Graphs: Methods and Applications (2020)
11. Farrant, K., Uddin, L.Q.: Atypical developmental of dorsal and ventral attention networks in autism. Developmental science **19**(4), 550–563 (2016)
12. Goldani, A.A., Downs, S.R., Widjaja, F., Lawton, B., Hendren, R.L.: Biomarkers in autism. Frontiers in psychiatry **5**, 100 (2014)
13. Gurbuz, M.B., Rekik, I.: Deep graph normalizer: A geometric deep learning approach for estimating connectional brain templates. In: International conference on medical image computing and computer-assisted intervention. pp. 155–165. Springer (2020)
14. Huang, Y., Chen, G., Wen, X., Shen, D.: Multi-modal brain graph learning of shared-specific features for schizophrenia diagnosis. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 1522–1525 (2024)
15. Kan, X., Cui, H., Lukemire, J., Guo, Y., Yang, C.: Fbnetgen: Task-aware gnn-based fmri analysis via functional brain network generation. In: International conference on medical imaging with deep learning. pp. 618–637. PMLR (2022)
16. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. Advances in Neural Information Processing Systems **35**, 25586–25599 (2022)
17. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. In: Advances in Neural Information Processing Systems. vol. 35, pp. 25586–25599 (2022)
18. Kawahara, J., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G.: Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage **146**, 1038–1049 (2017)
19. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L.H., Ventola, P., Duncan, J.S.: Braingnn: Interpretable brain graph neural network for fmri analysis. Medical Image Analysis **74**, 102233 (2021)
20. Liao, X., Vasilakos, A.V., He, Y.: Small-world human brain networks: perspectives and challenges. Neuroscience & Biobehavioral Reviews **77**, 286–300 (2017)
21. Meunier, D., Lambiotte, R., Fornito, A., Ersche, K., Bullmore, E.T.: Hierarchical modularity in human brain functional networks. Frontiers in neuroinformatics **3**, 571 (2009)
22. Padmanabhan, A., Lynch, C.J., Schaer, M., Menon, V.: The default mode network in autism. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **2**(6), 476–486 (2017)
23. Peng, L., Wang, N., Xu, J., Zhu, X., Li, X.: Gate: Graph cca for temporal self-supervised learning for label-efficient fmri analysis. IEEE Transactions on Medical Imaging **42**(2), 391–402 (2022)
24. Price, T., Wee, C.Y., Gao, W., Shen, D.: Multiple-network classification of childhood autism using functional connectivity dynamics. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 177–184. Springer (2014)

25. Rosazza, C., Minati, L.: Resting-state brain networks: literature review and clinical applications. Neurological sciences **32**(5), 773–785 (2011)
26. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. Neuroimage **52**(3), 1059–1069 (2010)
27. Shi, F., Wang, L., Peng, Z., Wee, C.Y., Shen, D.: Altered modular organization of structural cortical networks in children with autism. PloS one **8**(5), e63131 (2013)
28. Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., et al.: Correspondence of the brain's functional architecture during activation and rest. Proceedings of the national academy of sciences **106**(31), 13040–13045 (2009)
29. Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., et al.: Functional connectomics from resting-state fmri. Trends in cognitive sciences **17**(12), 666–682 (2013)
30. Sporns, O.: Network attributes for segregation and integration in the human brain. Current opinion in neurobiology **23**(2), 162–171 (2013)
31. Sporns, O., Honey, C.J., Kötter, R.: Identification and classification of hubs in brain networks. PloS one **2**(10), e1049 (2007)
32. van den Heuvel, M.P., Hulshoff Pol, H.E.: Exploring the brain network: A review on resting-state fmri functional connectivity. European Neuropsychopharmacology **20**(8), 519–534 (2010)
33. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph attention networks. In: International Conference on Learning Representations (2018)
34. Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology (2011)